DOCUMENT RESUME

ED 317 569 TM 014 608

AUTHOR Lawson, Stephen
TITLE The Importance of Computer Programming Skills to Educational Researchers.
PUB DATE Jan 90
NOTE 24p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 25-27, 1990).
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Software; *Computer Uses in Education; Educational Research; *Educational Researchers; *Factor Analysis; Literature Reviews; Microcomputers; *Programing; Programing Languages; *Research Methodology; *Statistical Analysis
IDENTIFIERS CENFAC Computer Program

ABSTRACT

The use of the modern computer has revolutionized the field of educational research. Software packages are currently available that allow almost anyone to analyze data efficiently and rapidly. Yet, caution must temper the widespread acceptance and use of these programs. It is recommended that the researcher not rely solely on the use of "canned" software program packages for analyzing data. It may also be important to learn to use a computer language, preferably FORTRAN or BASIC, to augment research skills and to create the opportunity for conducting analyses that are not automated in statistical packages. Several of these languages can be used to write programs that perform sophisticated analyses even on microcomputers. A computer program written in FORTRAN-77 is presented as an example. This program, CENFAC, extracts factor pattern coefficients from a correlation matrix utilizing the centroid method of factor extraction. Although this method of factor analysis has been largely abandoned in favor of others, it is useful for teaching students the concepts underlying factor analysis and other basic concepts of statistics. Tabulated output from the CENFAC program is appended. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED317569

computer.wp 1/22/90

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEPHEN LAWSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Importance of Computer Programming Skills to Educational Researchers

Stephen Lawson

University of New Orleans 70148

Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX, January 25, 1990. A copy of the program, "CENFAC", and supporting documentation may be obtained by contacting: Stephen Lawson, University of New Orleans, Lakefront Campus, ED #246, New Orleans, LA 70148

TM014608

Abstract

The use of the modern computer has revolutionized the field of educational research. Software packages are currently available that allow most anyone to analyze data efficiently and rapidly. Yet, caution must temper the widespread acceptance and use of these programs. It is recommended that the researcher not rely solely on the use of "canned" software program packages for analyzing data. It may also be important to learn to use a computer language, preferably FORTRAN or BASIC, to augment research skills and to create the opportunity for conducting analyses that are not automated in statistical packages. Several of these languages can be used to write programs that perform sophisticated analyses even on microcomputers. A computer program written in FORTRAN-77 is presented as an example.

Computer usage today has clearly revolutionized the field of behavioral research. Both the large, facility-based mainframe and the home personal computer (PC) have had a significant impact on the manner in which research is conducted. With statistical software programs available to perform so many data manipulations and analyses, the behavioral scientist has entered a new era of research.

The modern computer has provided the researcher with the power and the flexibility to conduct "more exhaustive, accurate, and affordable numerical analysis and statistical research" (Goldstein, 1989, p. 94). Analyses involving extensive computations that would have been unmanageable for the average individual to calculate by hand can be accurately performed today on a computer in a relatively short period of time. Consider the use of factor analysis. Although this technique was originally conceptualized prior to 1905, until fairly recently the method was rarely used due to the inordinate amount of calculations involved in extracting and rotating factor pattern coefficients. Today, however, research designs employing factor analysis as either the principal or secondary analysis are quite common, thanks largely to the efficient capabilities of the computer (Kerlinger, 1986).

The computer's profound influence on the field of behavioral research is also evident in the use of more recently developed statistical techniques. Intensive calculations that previously could only be approximated or imagined can be accurately performed

today on the computer. For example, the technique of bootstrapping (Diaconis & Efron, 1983) relies on the computer's number crunching capabilities to sample hundreds of artificial data sets from an original data set in order to test the reliability of statistical inference. Another technique, that of log-linear analysis, requires the computer to construct models of linear relationships, expressed in natural logarithms, among a group of variables (Rice, in press). So crucial is use of a computer to performing this analysis that Kerlinger (1986) stated, "Log-linear analysis ... is unthinkable without the computer. It is even unlikely that Goodman and others would have developed the theory and the extensive methodology without knowing that its tedious calculations could be done by computer" (p. 631).

Perhaps the most evident impact that the computer has had on field of behavioral research is that it has allowed for an increasing number of novice researchers to independently perform complex statistical analyses. The widespread availability of **user-friendly** statistical software packages has equipped all individuals, from the undergraduate student, to the marketing executive, to the department store clerk, with the ability to analyze their own data. Kerlinger (1986) may have appropriately summed up the situation by declaring, "Mathematical and statistical ignorance, too much a thing of the past, is no longer possible." (p. 631).

Statistical analyses can be accomplished on the computer

through any of a variety of software packages. The most widely used and preferred programs (Shavelson, 1988) have traditionally been BMDP (Biomedical Computer Programs) (Dixon, 1983), SPSS (Statistical Package for the Social Sciences) (Norusis, 1988), and SAS (Statistical Analysis Systems) (SAS Institute, Inc., 1982) . These packages were developed in late 1960s and early 1970s, originally for use on a mainframe computer. However, with the growth in popularity of the microcomputer during the 1980s, modified versions of these programs were quickly introduced and marketed to an growing population of home computer users.

Today, over 200 commercially produced statistical packages are available for use on the personal computer alone, along with several hundred shareware programs (Raskin, 1989). Some of these programs provide the researcher with basic statistical analyses (e.g., correlations, analyses of variance, and cross tabulations). Add-on packages expand the capabilities of some programs by providing the user with the option of selecting more advanced techniques such as factor analysis and discriminant analysis. Still other statistical packages constitute specialty items that are designed to appeal to a particular market. One such program "attempts to make statistics a part of everyday business by combining statistical functions with, of all things, address book, calculator, and calendar accessories" (Raskin, 1989, p.105).

The influence of these statistical software programs on the field of behavioral research has obviously been positive. Yet,

some caution must be exercised in using these packages, as noted by various researchers and authors. As Kerlinger (1986, p. 633) suggests, "The principal and underlying danger is that we slide too easily into letting the programs do our real work for us, and more important..., we reinforce the tendency of ourselves and our students to escape the necessity of knowing as clearly as possible what we are doing."

The "user-friendliness" of these programs allows the inexperienced researcher to readily arrive at a conclusion regarding hypotheses without ever having a clear understanding of the methodology involved in the analysis. By selecting any of a variety of menu options, individuals can analyze their data, receive printed results, and never know whether the selected statistical analysis was appropriate or not. Consequently, the output provided by many of these programs is often "misunderstood or ... misinterpreted by many researchers" (Barcikowski, 1983, p. xi).

Thus, Hinkle, Wiersma, and Jurs (1988) have rightfully warned:

> Although these packages relieve the researcher of the computational burden formerly associated with statistics, they do not relieve them of the responsibility for appropriate application of statistical procedures and proper interpretation of the results. This is especially true in light of the assumptions underlying various statistical procedures. (p. 632)

Researchers should also be cautioned not to accept the worth of a statistical program based on faith alone. Simply because a software package has been copyrighted and sold does not necessarily mean that it performs all analyses without fault. Gorsuch (1974) observed that "(w)hen a program does run and gives some output, there is a tendency to automatically assume the answers are correct. Unfortunately, this is often not the case" (p. 108). Bugs or errors in software programs are extremely common. As Raskin (1989, p. 104) notes, "Statistical packages, especially PC-based, menu-driven ones can look deceptively simple.... These packages can give the appearance of doing meaningful work when the results are nonsense." In short, all programs do not produce the same results.

For example, after reviewing the SAS (1982 version), SPSS (Version H), and BMDP (1981 version) programs, MacCallum (1983) strongly advised against the general use of SPSS for factor analysis. Citing the unavailability of proper eigenvalues produced by the principal axis option, along with other reasons, MacCallum concluded that the SPSS package was significantly inferior to the other two packages when used for factor analysis. Similarly, Huberty and Wisenbaker (in press) caution against the use of the SPSS-X discriminant analysis procedure when a quadratic or nonpooled classification rule is employed.

In another instance of software errors, Raskin (1989) reported that while preparing for a review of various statistical software

packages, an ANOVA procedure was performed on a data set using an available program. The results obtained from the program were sent to a group of reviewers who had no knowledge of the software package that was used to analyze the data. None of the reviewers could match the results. Only then was an error in the algorithm utilized by the program discovered.

While many software development companies are continuously upgrading their programs, users should nonetheless recognize that the version which they are using may not be 100 percent accurate. Kerlinger (1986) therefore emphasized that "package programs must be used with great care and that any particular program be checked for accuracy and adequate documentation" (p. 633).

Researchers should finally be cognizant of the fact that no one statistical software package will do everything. Computer programs, as with any commercially marketed product, are designed with the basic purpose of being sold and earning a profit. The most successful product is generally the one that appeals to the widest segment of the market. Certain statistical tests and procedures, therefore, are often not included within software packages simply because they do not appeal to enough users to generate interest in the package.

In like manner, software programs are usually not able to represent more recently developed statistical procedures. Research methodology is constantly undergoing development and extension. Yet, as newer techniques are accepted and utilized by the field,

years may pass before the methodology is programmed, documented, and corrected so that it may be included for use within a commercial statistical package. Thus, journals such as Educational and Psychological Measurement regularly publish articles describing more specialized computer programs that can supplement the routines available in packages.

Regarding the previously mentioned cautions against blithely using commercial statistical software programs, Kerlinger (1986) provides the following suggestions. First, the behavioral researcher must have enough methodological competence to do most things independently. Complete reliance on software programs, computer specialists, or statistical consultants for performing data analysis is inexcusable. In such instances, the researcher often allows himself or herself to rush into data collection and analysis without having thoroughly studied the problem. Simply put, "(t)he behavioral scientist... has to know enough of statistical analysis and computer technology to be able to use technicians as resources rather than as preceptors" (Kerlinger, 1986, p. 635).

Second, it is recommended that the researcher learn at least one computer language, preferably FORTRAN or BASIC. FORTRAN (Formula Translation) was originally developed in the late 1950s so that engineers could easily write and understand their own programs without having to use complex machine language. While recently criticized as a "dead language" (cf. Krus & Lu, 1987),

FORTRAN has never-the-less remained the dominant programming language for the behavioral researcher due to its strength in manipulating and analyzing data (Cole, 1987). BASIC is the programming language utilized by most personal computers. Designed so that most college students could easily learn to use it, BASIC is also powerful enough for most applications in the behavioral sciences.

Heuristic Example of a Useful Program

This paper presents, as an example of the use of computer languages to write helpful programs for analyzing educational research data, a computer program written in FORTRAN-77. Program CENFAC extracts factor pattern coefficients from a correlation matrix utilizing the centroid method of factor extraction (cf. Thurstone, 1947). This method of factor analysis was at one time quite popular, since it can be performed by hand using only pencil and paper. With the advent of the computer it is rarely used anymore, if at all, in favor of other methods. However, such a program can be very useful in teaching students about the concepts underlying factor analysis. This use highlights another application of programming knowledge--writing software that helps students to understand statistical concepts or procedures.

Centroid Method

The centroid method provides the researcher with excellent experience in learning how factors are extracted from a matrix of correlations. The steps are as follows. First, each column in

the matrix is summed. These figures are added together to form a grand sum. Each column sum is then divided by the square root of the grand sum, providing the pattern coefficients for the first factor. As with principal components, this factor accounts for the largest amount of variance in the matrix.

Second, a matrix of cross products is formed using the pattern coefficients for the first factor. The matrix of cross products is subtracted from the original correlation matrix to form a matrix of residuals. These columns are then summed, excluding the diagonal entry. All entries in the column with the largest negative sum, and the entries in the corresponding row, are then reflected or reversed. Columns are summed again, without the diagonal entry, and the process of reflection is continued until all column sums, each excluding the diagonal entries, are positive.

The second factor pattern coefficients are then extracted from this matrix in the same manner as the first factor. The entire process is continued, with each factor accounting for decreasing amounts of variance, until the majority of variance in the matrix has been accounted for by the extracted factors.

Partial output from program CENFAC can be found in the accompanying appendix. The initial matrix for this demonstration consisted of the correlations among responses of special education administrators, teachers, and parents to 15 survey items (Lawson, 1988). Steps for extracting factors from the matrix using the centroid method are listed in the printout, along with each of the

resulting matrices. In this manner, users of the program can closely follow each step of the procedure.

Input

Required input includes a parameter card declaring data format and the number of variables in the original correlation matrix. Data are read from a file named, "FOR025.DAT".

Output

The output of the program is contained in the file, "FOR026.DAT". Factors are extracted for eigenvalues greater than or equal to 1. Program CENFAC provides output from each step of the centroid process for the user, along with documentation. All correlation, residual, and reflected matrices are provided so that the user can gain a greater understanding of how factors are extracted from a matrix of correlations.

Summary

The use of the modern computer has revolutionized the field of educational research. Software packages are currently available that allow most anyone to analyze data efficiently and rapidly. Yet, caution must temper the widespread acceptance and use of these programs. It is recommended that the researcher not rely solely on the use of "canned" software program packages for analyzing data. It may also be important to learn to use a computer language, preferably FORTRAN or BASIC, to augment research skills and to create the opportunity for conducting analyses that are not automated in statistical packages. Several of these languages can

be used to write programs that perform sophisticated analyses even on microcomputers. A computer program written in FORTRAN-77 was presented as an example. Although some programs are written to implement analyses not available in software packages, program CENFAC illustrates the use of programs to teach students basic statistics concepts.

References

Barcikowski, R.S. (Ed.) (1983). Computer packages and research design (Vol. 3: SPSS and SPSSX). Lanham, MD: University Press of America, Inc.

Cole, J.W.P. (1987). ANSI FORTRAN 77: A Structured Problem-Solving Approach. Dubuque, IA: Wm. C. Brown Publishers.

Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. Scientific American, 248(5), 116-130.

Dixon, W.J. (1981). BMDP Statistical Software-1981. Berkeley, CA: University of California Press.

Goldstein, R. (1989, March 14). Understanding numerical analysis: Working with the odds. PC Magazine, pp. 94-99.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (1988). Applying Statistics for the Behavioral Sciences (2nd ed.). Boston, MA: Houghton Mifflin Co.

Huberty, C., & Wisenbaker, J. (in press). Discriminant analysis: Potential improvements in typical practice. In B. Thompson (Ed.), Advances in social science methodology (Vol. 2). Greenwich, CT: JAI Press.

Kerlinger, F.N. (1986). Foundations of behavioral research (3rd ed.). New York: Holt, Rinehart and Winston, Inc.

Krus, D.J., & Lu, M. (1987). Bias in computer languages comparisons: A FORTRAN phobic cabal? Educational and Psychological Measurement, 47, 643-647.

Lawson, S. (1988). The validation of transition concepts for young

adults with severe handicaps in the Jefferson Parish (Louisiana) public school system. Unpublished master's thesis, University of New Orleans, New Orleans, LA.

MacCallum, R. (1983). A comparison of factor analysis programs in SPSS, BDMP, and SAS. Psychometrika, 48(2), 223-231.

Norusis, M. (1988). SPSS-X User's Guide (3rd ed.). Chicago, IL: SPSS Inc.

Raskin, R. (1989, March 14). Statistical software for the PC: Testing for significance. PC Magazine, pp. 103-116.

Rice, J. (in press). Loglinear analysis: Analysis of categorical variables in the logit setting. In B. Thompson (Ed.), Advances in social science methodology (Vol. 2). Greenwich, CT: JAI Press.

SAS Institute Inc. SAS User's Guide: Statistics (1982 ed.) Cary, NC: SAS Institute Inc.

Shavelson, R.J. (1988). Statistical reasoning for the behavioral sciences (2nd ed.). Boston, MA: Allyn & Bacon Inc.

Thurstone, L.L. (1947). Multiple-Factor Analysis. Chicago, IL: University of Chicago Press.

Appendix

Partial Output from Program CENFAC

****

ORIGINAL CORRELATION MATRIX

****

1.000	0.086	0.505	0.312	0.246	0.382	0.201	0.315	0.211	0.162	0.326	0.364	0.449	0.494	0.383
0.086	1.000	0.356	0.300	0.387	0.272	0.279	-0.007	0.065	0.073	0.147	0.013	0.114	0.021	-0.026
0.505	0.356	1.000	0.573	0.603	0.668	0.335	0.397	0.242	0.229	0.339	0.334	0.507	0.372	0.332
0.312	0.300	0.573	1.000	0.485	0.462	0.273	0.241	0.270	0.286	0.347	0.368	0.420	0.282	0.304
0.246	0.387	0.603	0.485	1.000	0.549	0.398	0.260	0.095	0.114	0.363	0.213	0.403	0.226	0.247
0.382	0.272	0.668	0.462	0.549	1.000	0.347	0.377	0.264	0.246	0.427	0.364	0.447	0.320	0.496
0.201	0.279	0.335	0.273	0.398	0.347	1.000	0.290	0.171	0.181	0.371	0.303	0.351	0.181	0.273
0.315	-0.007	0.397	0.241	0.260	0.377	0.290	1.000	0.383	0.418	0.377	0.334	0.334	0.337	0.245
0.211	0.065	0.242	0.270	0.095	0.264	0.171	0.383	1.000	0.501	0.262	0.218	0.207	0.363	0.293
0.162	0.073	0.229	0.286	0.114	0.246	0.181	0.418	0.501	1.000	0.400	0.419	0.311	0.210	0.244
0.326	0.147	0.339	0.347	0.363	0.427	0.371	0.377	0.262	0.400	1.000	0.543	0.495	0.379	0.420
0.364	0.013	0.334	0.368	0.213	0.364	0.303	0.334	0.218	0.419	0.543	1.000	0.510	0.413	0.481
0.449	0.114	0.507	0.420	0.403	0.447	0.351	0.334	0.207	0.311	0.495	0.510	1.000	0.466	0.564
0.494	0.021	0.372	0.282	0.226	0.320	0.181	0.337	0.363	0.210	0.379	0.413	0.466	1.000	0.547
0.383	-0.026	0.332	0.304	0.247	0.496	0.273	0.245	0.293	0.244	0.420	0.481	0.564	0.547	1.000

Note that this is the original correlation matrix as supplied by the user.

****

FACTOR MATRIX A

****

1.000	0.086	0.505	0.312	0.246	0.382	0.201	0.315	0.211	0.162	0.326	0.364	0.449	0.494	0.383
0.086	1.000	0.356	0.300	0.387	0.272	0.279	-0.007	0.065	0.073	0.147	0.013	0.114	0.021	-0.026
0.505	0.356	1.000	0.573	0.603	0.668	0.335	0.397	0.242	0.229	0.339	0.334	0.507	0.372	0.332
0.312	0.300	0.573	1.000	0.485	0.462	0.273	0.241	0.270	0.286	0.347	0.368	0.420	0.282	0.304
0.246	0.387	0.603	0.485	1.000	0.549	0.398	0.260	0.095	0.114	0.363	0.213	0.403	0.226	0.247
0.382	0.272	0.668	0.462	0.549	1.000	0.347	0.377	0.264	0.246	0.427	0.364	0.447	0.320	0.496
0.201	0.279	0.335	0.273	0.398	0.347	1.000	0.290	0.171	0.181	0.371	0.303	0.351	0.181	0.273
0.315	-0.007	0.397	0.241	0.260	0.377	0.290	1.000	0.383	0.418	0.377	0.334	0.334	0.337	0.245
0.211	0.065	0.242	0.270	0.095	0.264	0.171	0.383	1.000	0.501	0.262	0.218	0.207	0.363	0.293
0.162	0.073	0.229	0.286	0.114	0.246	0.181	0.418	0.501	1.000	0.400	0.419	0.311	0.210	0.244
0.326	0.147	0.339	0.347	0.363	0.427	0.371	0.377	0.262	0.400	1.000	0.543	0.495	0.379	0.420
0.364	0.013	0.334	0.368	0.213	0.364	0.303	0.334	0.218	0.419	0.543	1.000	0.510	0.413	0.481
0.449	0.114	0.507	0.420	0.403	0.447	0.351	0.334	0.207	0.311	0.495	0.510	1.000	0.466	0.564
0.494	0.021	0.372	0.282	0.226	0.320	0.181	0.337	0.363	0.210	0.379	0.413	0.466	1.000	0.547
0.383	-0.026	0.332	0.304	0.247	0.496	0.273	0.245	0.293	0.244	0.420	0.481	0.564	0.547	1.000
5.436	3.080	6.792	5.923	5.589	6.621	4.954	5.301	4.545	4.794	6.196	5.877	6.578	5.611	5.803
.59632	.33787	.74507	.64974	.61310	.72631	.54344	.58151	.49858	.52589	.67969	.64470	.72159	.61552	.63658

STEP #1: Sum each of the columns in the correlation matrix (without the diagonal entries).
The column sums (without the diagonal entries) are contained in the second to last row of the matrix.

STEP #2: Sum each of the column sums (without the diagonal entries).
The sum of the column sums = 83.100

STEP #3: Calculate the square root of the sum of sums.
The square root of the sum of sums = 9.115921

STEP #4: Divide the square root of the sum of sums into each of the column sums (without the diagonal entries).
The resulting figures are the factor pattern coefficients of Factor #1 and are contained in the last row of the matrix.

Note that these coefficients can be found in the Factor Pattern Matrix located on the last page of the printout.

****

MATRIX OF CROSS PRODUCTS (#1)

****

0.356	0.201	0.444	0.387	0.366	0.433	0.324	0.347	0.297	0.314	0.405	0.384	0.430	0.367	0.380
0.201	0.114	0.252	0.220	0.207	0.245	0.184	0.196	0.168	0.178	0.230	0.218	0.244	0.208	0.215
0.444	0.252	0.555	0.484	0.457	0.541	0.405	0.433	0.371	0.392	0.506	0.480	0.538	0.459	0.474
0.387	0.220	0.484	0.422	0.398	0.472	0.353	0.378	0.324	0.342	0.442	0.419	0.469	0.400	0.414
0.366	0.207	0.457	0.398	0.376	0.445	0.333	0.357	0.306	0.322	0.417	0.395	0.442	0.377	0.390
0.433	0.245	0.541	0.472	0.445	0.528	0.395	0.422	0.362	0.382	0.494	0.468	0.524	0.447	0.462
0.324	0.184	0.405	0.353	0.333	0.395	0.295	0.316	0.271	0.286	0.369	0.350	0.392	0.334	0.346
0.347	0.196	0.433	0.378	0.357	0.422	0.316	0.338	0.290	0.306	0.395	0.375	0.420	0.358	0.370
0.297	0.168	0.371	0.324	0.306	0.362	0.271	0.290	0.249	0.262	0.339	0.321	0.360	0.307	0.317
0.314	0.178	0.392	0.342	0.322	0.382	0.286	0.306	0.262	0.277	0.357	0.339	0.379	0.324	0.335
0.405	0.230	0.506	0.442	0.417	0.494	0.369	0.395	0.339	0.357	0.462	0.438	0.490	0.418	0.433
0.384	0.218	0.480	0.419	0.395	0.468	0.350	0.375	0.321	0.339	0.438	0.416	0.465	0.397	0.410
0.430	0.244	0.538	0.469	0.442	0.524	0.392	0.420	0.360	0.379	0.490	0.465	0.521	0.444	0.459
0.367	0.208	0.459	0.400	0.377	0.447	0.334	0.358	0.307	0.324	0.418	0.397	0.444	0.379	0.392
0.380	0.215	0.474	0.414	0.390	0.462	0.346	0.370	0.317	0.335	0.433	0.410	0.459	0.392	0.405

STEP #5: Create a matrix of cross products using the Factor # 1 pattern coefficients.

****

MATRIX OF RESIDUALS (#1)

****

0.644	-0.115	0.061	-0.075	-0.120	-0.051	-0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
-0.115	0.886	0.104	0.080	0.180	0.027	0.095	-0.203	-0.103	-0.105	-0.083	-0.205	-0.130	-0.187	-0.241
0.061	0.104	0.445	0.089	0.146	0.127	-0.070	-0.036	-0.129	-0.163	-0.167	-0.146	-0.031	-0.087	-0.142
-0.075	0.080	0.089	0.578	0.087	-0.010	-0.080	-0.137	-0.054	-0.056	-0.095	-0.051	-0.049	-0.118	-0.110
-0.120	0.180	0.146	0.087	0.624	0.104	0.065	-0.097	-0.211	-0.208	-0.054	-0.182	-0.039	-0.151	-0.143
-0.051	0.027	0.127	-0.010	0.104	0.472	-0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
-0.123	0.095	-0.070	-0.080	0.065	-0.048	0.705	-0.026	-0.100	-0.105	0.002	-0.047	-0.041	-0.153	-0.073
-0.032	-0.203	-0.036	-0.137	-0.097	-0.045	-0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	-0.103	-0.129	-0.054	-0.211	-0.098	-0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	-0.105	-0.163	-0.056	-0.208	-0.136	-0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	-0.083	-0.167	-0.095	-0.054	-0.067	0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	-0.205	-0.146	-0.051	-0.182	-0.104	-0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	-0.130	-0.031	-0.049	-0.039	-0.077	-0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	-0.187	-0.087	-0.118	-0.151	-0.127	-0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	-0.241	-0.142	-0.110	-0.143	0.034	-0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.644	-0.886	-0.445	-0.578	-0.624	-0.472	-0.705	-0.662	-0.751	-0.723	-0.538	-0.584	-0.479	-0.621	-0.595

STEP #6: Subtract the matrix of cross products from the original matrix to create a matrix of residuals.

The second to last row of the matrix contains the column sums with the diagonal entry.
Note that these column sums equal zero within rounding error.

The last row of the matrix contains the column sums without the diagonal entry.
Note that these column sums are equal to the diagonal entries within rounding error.

STEP #7: Locate the largest negative column sum (without the diagonal entry). Reflect (reverse) the signs of all entries in that column and in the corresponding row.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 1
Note that Column # 2 and Row # 2 have been reflected.

****

0.644	0.115	0.061	-0.075	-0.120	-0.051	-0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	-0.104	-0.080	-0.180	-0.027	-0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
0.061	-0.104	0.445	0.089	0.146	0.127	-0.070	-0.036	-0.129	-0.163	-0.167	-0.146	-0.031	-0.087	-0.142
-0.075	-0.080	0.089	0.578	0.087	-0.010	-0.080	-0.137	-0.054	-0.056	-0.095	-0.051	-0.049	-0.118	-0.110
-0.120	-0.180	0.146	0.087	0.624	0.104	0.065	-0.097	-0.211	-0.208	-0.054	-0.182	-0.039	-0.151	-0.143
-0.051	-0.027	0.127	-0.010	0.104	0.472	-0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
-0.123	-0.095	-0.070	-0.080	0.065	-0.048	0.705	-0.026	-0.100	-0.105	0.002	-0.047	-0.041	-0.153	-0.073
-0.032	0.203	-0.036	-0.137	-0.097	-0.045	-0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	-0.129	-0.054	-0.211	-0.098	-0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	-0.163	-0.056	-0.208	-0.136	-0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	-0.167	-0.095	-0.054	-0.067	0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	-0.146	-0.051	-0.182	-0.104	-0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	-0.031	-0.049	-0.039	-0.077	-0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	-0.087	-0.118	-0.151	-0.127	-0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	-0.142	-0.110	-0.143	0.034	-0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
-0.413	0.886	-0.653	-0.739	-0.984	-0.526	-0.895	-0.255	-0.545	-0.514	-0.373	-0.175	-0.220	-0.247	-0.113

The last row of the matrix contains the column sums without the diagonal entry.

STEP #8: Repeat STEP #7 until all column sums (without the diagonal entries) are positive.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 2
Note that Column # 5 and Row # 5 have been reflected.

****

0.644	0.115	0.061	-0.075	0.120	-0.051	-0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	-0.104	-0.080	0.180	-0.027	-0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
0.061	-0.104	0.445	0.089	-0.146	0.127	-0.070	-0.036	-0.129	-0.163	-0.167	-0.146	-0.031	-0.087	-0.142
-0.075	-0.080	0.089	0.578	-0.087	-0.010	-0.080	-0.137	-0.054	-0.056	-0.095	-0.051	-0.049	-0.118	-0.110
0.120	0.180	-0.146	-0.087	0.624	-0.104	-0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
-0.051	-0.027	0.127	-0.010	-0.104	0.472	-0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
-0.123	-0.095	-0.070	-0.080	-0.065	-0.048	0.705	-0.026	-0.100	-0.105	0.002	-0.047	-0.041	-0.153	-0.073
-0.032	0.203	-0.036	-0.137	0.097	-0.045	-0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	-0.129	-0.054	0.211	-0.098	-0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	-0.163	-0.056	0.208	-0.136	-0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	-0.167	-0.095	0.054	-0.067	0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	-0.146	-0.051	0.182	-0.104	-0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	-0.031	-0.049	0.039	-0.077	-0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	-0.087	-0.118	0.151	-0.127	-0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	-0.142	-0.110	0.143	0.034	-0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
-0.174	1.246	-0.946	-0.912	0.984	-0.733	-1.025	-0.062	-0.123	-0.097	-0.265	0.190	-0.141	0.056	0.174

The last row of the matrix contains the column sums without the diagonal entry.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 3
Note that Column # 7 and Row # 7 have been reflected.

****

0.644	0.115	0.061	-0.075	0.120	-0.051	0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	-0.104	-0.080	0.180	-0.027	0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
0.061	-0.104	0.445	0.089	-0.146	0.127	0.070	-0.036	-0.129	-0.163	-0.167	-0.146	-0.031	-0.087	-0.142
-0.075	-0.080	0.089	0.578	-0.087	-0.010	0.080	-0.137	-0.054	-0.056	-0.095	-0.051	-0.049	-0.118	-0.110
0.120	0.180	-0.146	-0.087	0.624	-0.104	0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
-0.051	-0.027	0.127	-0.010	-0.104	0.472	0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
0.123	0.095	0.070	0.080	0.065	0.048	0.705	0.026	0.100	0.105	-0.002	0.047	0.041	0.153	0.073
-0.032	0.203	-0.036	-0.137	0.097	-0.045	0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	-0.129	-0.054	0.211	-0.098	0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	-0.163	-0.056	0.208	-0.136	0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	-0.167	-0.095	0.054	-0.067	-0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	-0.146	-0.051	0.182	-0.104	0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	-0.031	-0.049	0.039	-0.077	0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	-0.087	-0.118	0.151	-0.127	0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	-0.142	-0.110	0.143	0.034	0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
0.072	1.436	-0.806	-0.752	1.113	-0.638	1.025	-0.010	0.077	0.112	-0.269	0.285	-0.059	0.363	0.320

The last row of the matrix contains the column sums without the diagonal entry.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 4
Note that Column # 3 and Row # 3 have been reflected.

****

0.644	0.115	-0.061	-0.075	0.120	-0.051	0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	0.104	-0.080	0.180	-0.027	0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
-0.061	0.104	0.445	-0.089	0.146	-0.127	-0.070	0.036	0.129	0.163	0.167	0.146	0.031	0.087	0.142
-0.075	-0.080	-0.089	0.578	-0.087	-0.010	0.080	-0.137	-0.054	-0.056	-0.095	-0.051	-0.049	-0.118	-0.110
0.120	0.180	0.146	-0.087	0.624	-0.104	0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
-0.051	-0.027	-0.127	-0.010	-0.104	0.472	0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
0.123	0.095	-0.070	0.080	0.065	0.048	0.705	0.026	0.100	0.105	-0.002	0.047	0.041	0.153	0.073
-0.032	0.203	0.036	-0.137	0.097	-0.045	0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	0.129	-0.054	0.211	-0.098	0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	0.163	-0.056	0.208	-0.136	0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	0.167	-0.095	0.054	-0.067	-0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	0.146	-0.051	0.182	-0.104	0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	0.031	-0.049	0.039	-0.077	0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	0.087	-0.118	0.151	-0.127	0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	0.142	-0.110	0.143	0.034	0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
-0.050	1.645	0.806	-0.930	1.406	-0.891	0.885	0.063	0.336	0.438	0.066	0.577	0.003	0.536	0.604

The last row of the matrix contains the column sums without the diagonal entry.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 5
Note that Column # 4 and Row # 4 have been reflected

****

0.644	0.115	-0.061	0.075	0.120	-0.051	0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	0.104	0.080	0.180	-0.027	0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
-0.061	0.104	0.445	0.089	0.146	-0.127	-0.070	0.036	0.129	0.163	0.167	0.146	0.031	0.087	0.142
0.075	0.080	0.089	0.578	0.087	0.010	-0.080	0.137	0.054	0.056	0.095	0.051	0.049	0.118	0.110
0.120	0.180	0.146	0.087	0.624	-0.104	0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
-0.051	-0.027	-0.127	0.010	-0.104	0.472	0.048	-0.045	-0.098	-0.136	-0.067	-0.104	-0.077	-0.127	0.034
0.123	0.095	-0.070	-0.080	0.065	0.048	0.705	0.026	0.100	0.105	-0.002	0.047	0.041	0.153	0.073
-0.032	0.203	0.036	0.137	0.097	-0.045	0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	0.129	0.054	0.211	-0.098	0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	0.163	0.056	0.208	-0.136	0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	0.167	0.095	0.054	-0.067	-0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	0.146	0.051	0.182	-0.104	0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	0.031	0.049	0.039	-0.077	0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	0.087	0.118	0.151	-0.127	0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	0.142	0.110	0.143	0.034	0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
0.101	1.806	0.984	0.930	1.579	-0.872	0.725	0.336	0.444	0.549	0.256	0.679	0.100	0.772	0.824

The last row of the matrix contains the column sums without the diagonal entry.

****

MATRIX OF RESIDUALS (#1), REFLECTION # 6
Note that Column # 6 and Row # 6 have been reflected

****

0.644	0.115	-0.061	0.075	0.120	0.051	0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	0.104	0.080	0.180	0.027	0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
-0.061	0.104	0.445	0.089	0.146	0.127	-0.070	0.036	0.129	0.163	0.167	0.146	0.031	0.087	0.142
0.075	0.080	0.089	0.578	0.087	-0.010	-0.080	0.137	0.054	0.056	0.095	0.051	0.049	0.118	0.110
0.120	0.180	0.146	0.087	0.624	0.104	0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
0.051	0.027	0.127	-0.010	0.104	0.472	-0.048	0.045	0.098	0.136	0.067	0.104	0.077	0.127	-0.034
0.123	0.095	-0.070	-0.080	0.065	-0.048	0.705	0.026	0.100	0.105	-0.002	0.047	0.041	0.153	0.073
-0.032	0.203	0.036	0.137	0.097	0.045	0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	0.129	0.054	0.211	0.098	0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	0.163	0.056	0.208	0.136	0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	0.167	0.095	0.054	0.067	-0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	0.146	0.051	0.182	0.104	0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	0.031	0.049	0.039	0.077	0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	0.087	0.118	0.151	0.127	0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	0.142	0.110	0.143	-0.034	0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
0.204	1.859	1.237	0.910	1.786	0.872	0.630	0.427	0.640	0.821	0.389	0.887	0.255	1.026	0.756

The last row of the matrix contains the column sums without the diagonal entry.

****

FACTOR MATRIX # 2

****

0.644	0.115	-0.061	0.075	0.120	0.051	0.123	-0.032	-0.086	-0.152	-0.079	-0.020	0.019	0.127	0.003
0.115	0.886	0.104	0.080	0.180	0.027	0.095	0.203	0.103	0.105	0.083	0.205	0.130	0.187	0.241
-0.061	0.104	0.445	0.089	0.146	0.127	-0.070	0.036	0.129	0.163	0.167	0.146	0.031	0.087	0.142
0.075	0.080	0.089	0.578	0.087	-0.010	-0.080	0.137	0.054	0.056	0.095	0.051	0.049	0.118	0.110
0.120	0.180	0.146	0.087	0.624	0.104	0.065	0.097	0.211	0.208	0.054	0.182	0.039	0.151	0.143
0.051	0.027	0.127	-0.010	0.104	0.472	-0.048	0.045	0.098	0.136	0.067	0.104	0.077	0.127	-0.034
0.123	0.095	-0.070	-0.080	0.065	-0.048	0.705	0.026	0.100	0.105	-0.002	0.047	0.041	0.153	0.073
-0.032	0.203	0.036	0.137	0.097	0.045	0.026	0.662	0.093	0.112	-0.018	-0.041	-0.086	-0.021	-0.125
-0.086	0.103	0.129	0.054	0.211	0.098	0.100	0.093	0.751	0.239	-0.077	-0.103	-0.153	0.056	-0.024
-0.152	0.105	0.163	0.056	0.208	0.136	0.105	0.112	0.239	0.723	0.043	0.080	-0.068	-0.114	-0.091
-0.079	0.083	0.167	0.095	0.054	0.067	-0.002	-0.018	-0.077	0.043	0.538	0.105	0.005	-0.039	-0.013
-0.020	0.205	0.146	0.051	0.182	0.104	0.047	-0.041	-0.103	0.080	0.105	0.584	0.045	0.016	0.071
0.019	0.130	0.031	0.049	0.039	0.077	0.041	-0.086	-0.153	-0.068	0.005	0.045	0.479	0.022	0.105
0.127	0.187	0.087	0.118	0.151	0.127	0.153	-0.021	0.056	-0.114	-0.039	0.016	0.022	0.621	0.155
0.003	0.241	0.142	0.110	0.143	-0.034	0.073	-0.125	-0.024	-0.091	-0.013	0.071	0.105	0.155	0.595
0.848	2.745	1.682	1.488	2.411	1.344	1.334	1.089	1.391	1.545	0.927	1.472	0.734	1.647	1.351

.18077 .58510 .35861 .31712 .51385 .28649 .28443 .23213 .29657 .32929 .19758 .31375 .15644 .35105 .28802

STEP #9: When all column sums (without the diagonal entries) are positive, repeat from STEP #1 to extract another factor.

The sum of the column sums = 22.007
The square root of the sum of sums = 4.691205.

The second to last row of the matrix contains the column sums without the diagonal entry.
The last row of the matrix contains the extracted factor pattern coefficients.

STEP #10: Continue with the above procedure until the majority of variance in the correlation matrix has been accounted for by the extracted factors.

****

FACTOR PATTERN MATRIX

****

0.5963	0.1808	0.1788	0.4203
0.3379	0.5851	0.2314	0.5101
0.7451	0.3586	0.0902	0.6919
0.6497	0.3171	0.2039	0.5643
0.6131	0.5139	0.1695	0.6687
0.7263	0.2865	0.1374	0.6285
0.5434	0.2844	0.2187	0.4241
0.5815	0.2321	0.3193	0.4940
0.4986	0.2966	0.5008	0.5873
0.5259	0.3293	0.4833	0.6185
0.6797	0.1976	0.1737	0.5312
0.6447	0.3137	0.1859	0.5486
0.7216	0.1564	0.2478	0.6066
0.6155	0.3511	0.1797	0.5344
0.6366	0.2880	0.3375	0.6021
5.692	1.650	1.089	
37.943	10.997	7.262	
37.943	48.941	56.203	

The first 3 columns contain the unrotated factor pattern coefficients.
The last column contains the communalities.

The first row beneath the matrix contains the eigenvalues for each extracted factor.

The second row beneath the matrix contains the variance explained by each factor.

The last row beneath the matrix contains the cumulative variance explained by each factor.